

Louis Miguel Bernal

+63 999 770 3445 | miguellouis.work@gmail.com | linkedin.com/in/louisbernal/ | github.com/LouisMiguelBernal | 1nonly.me

TECHNICAL SKILLS

Programming Languages & Tools: Python, SQL, Java, C++, JavaScript, Next.js, React, Docker, Linux, Git/GitHub, REST APIs

Data Engineering & AI Stack: AWS Bedrock, RAG, LLMs, Snowflake, dbt, Airflow, Azure, PyTorch, TensorFlow, Computer Vision, XGBoost, Scikit-learn, NLP

WORK EXPERIENCE

Freelance

Remote

AI Data Engineer | Python, Airflow, dbt, Snowflake, AWS Bedrock, RAG, SQL, Git

June 2026 – Present

- Built AWS Bedrock LLM services over the warehouse: document intelligence parsing loan files, appraisals and financials at 97% field accuracy, a confidence-scored entity-resolution matcher reconciling hundreds of properties against a 14K-row regulatory registry, and a RAG copilot on the Snowflake semantic layer, cutting manual entry 80%.
- Architected and own the Airflow DAG ingestion layer across 9 enterprise source systems, including Hypha, NetSuite and iLevel, merging into tiered Snowflake raw schemas with source and load-timestamp lineage columns and moving 2M+ records/day at 99.9% uptime behind keyset-paged loaders and loud-failure schema-drift guards.
- Engineered a multi-layer dbt transformation backbone on Snowflake (staging, intermediate, export and mart layers, 350+ models), standardizing 120+ tables behind reusable data-quality and performance-metric test macros and a severity-tagged validation catalogue, cutting model build and QC time 60% and warehouse compute spend 35%.
- Delivered the finance and accounting pipelines the business runs on, rebuilding a portfolio valuation model to a 99.7% identity match against the reporting tape, shipping general-ledger ingestion with automated reconciliation checks, and automating an underwriting-to-asset-system sync with read-compare-write logic that never overwrites a human entry.

Float Infinity

Sydney, Australia

Data Engineer, Full-Stack Software Engineer & Quant | Azure, Python, SQL, Airflow, dbt, XGBoost, Next.js

March 2026 – June 2026

- Architected an automated Azure telemetry and reporting platform ingesting real-time signals from Azure Resources, Policy, Patching, and Backup through REST API pipelines orchestrated on Airflow DAGs, cutting incident detection time by 50%, slashing reporting overhead by 70-80%, and sustaining 99.9% availability.
- Engineered production ELT on Azure, pairing Python/SQLAlchemy ingestion with modular, version-controlled dbt models using incremental materialization and unique-key deduplication, orchestrated end-to-end by Airflow DAGs across 4+ enterprise platforms, cutting data redundancy by 65%, tripling query performance, and lifting cloud governance visibility by 80%+.
- Built an ASX options trading analytics platform end to end: engineered a provider-agnostic market data ingestion service (Python, FastAPI) pulling OHLCV price history into PostgreSQL via idempotent upserts and watchlist sync, exposed through an authenticated REST API, and built the Next.js/React frontend stock list/detail pages with candlestick charting and watchlist management using optimistic UI updates backed by an integration test suite against a live PostgreSQL database and CI.
- Developed an end-to-end NLP/ML quantitative trading algorithm across 15+ instruments, training sentiment and macro-regime models on real-time market sentiment, macroeconomic indicators, institutional positioning flows, and options data. Engineered 50+ features across technical momentum, sentiment, and macroeconomic layers for 72% directional accuracy and a backtested Sharpe Ratio of 1.8.

PASIA – Procurement and Supply Institute of Asia

Makati, Metro Manila

Data Analyst | Python, SQL, Scikit-learn, Pandas, Power BI, Excel, Git

June 2025 – Aug 2025

- Automated large-scale data preprocessing pipelines using Python (pandas, scikit-learn) and machine learning–based imputation, processing over 1 million procurement and contract records, reducing manual effort by 85%.
- Optimized SQL ETL pipelines for automated ingestion and transformation of high-volume procurement data, ensuring 99% consistency across departments and daily data refreshes for real-time insights.
- Developed advanced Power BI and Excel visualizations of contractor procurement data, transforming 1M+ contract records into insights that surfaced high-value savings and efficiency gains.

Freelance

Remote

Full-Stack Web Developer & Data Analyst | JavaScript, React, Next.js, Excel, Python, Power BI

Sept 2022 – Feb 2026

- Designed and shipped 40+ responsive websites and web applications for small-business and academic clients using JavaScript, React, and Next.js, owning each project from requirements gathering through production deployment.
- Delivered Power BI and Excel dashboards for business and academic clients, turning raw data into decision-ready insights and cutting manual reporting time by ~40%.
- Engineered repeatable data-cleaning and EDA workflows in Python (Pandas) and Excel, standardizing ingestion across client datasets and reducing manual prep time by ~40%.

PROJECTS

Nexus: AI-Powered Institutional Multi-Engine Alpha & Risk System | Python, SQLAlchemy, LLMs

March 2026 – Present

- Built a low-latency multi-venue derivatives trading system ingesting full-depth order book data from Binance, OKX, and MEXC with 100ms–2s signal latency, implementing cross-venue book fusion.
- Engineered an institutional-grade risk and execution framework spanning 15+ venues/instruments, combining multi-model VaR, vol/correlation-adjusted Kelly sizing, and an event-driven circuit breaker (async pub/sub) reacting to risk events in <10ms.
- Designed risk intelligence pipelines with AI-powered monitoring (Gemma 4 e4b + FinBERT) for anomaly detection, exposure tracking, and adaptive risk throttling, achieving 99% detection accuracy/precision.

VisorAI: A YOLOv11-Powered Mobile Application Detection with Auditory Feedback | CV, Flutter *May 2025 – March 2026*

- Engineered a real-time object detection system using YOLOv11 for automated recognition of Philippine LTO road markings and traffic signage, achieving 0.87 mAP@0.5 across 12 marking classes.
- Architected a cross-platform Flutter app integrating the optimized CV inference pipeline with a text-to-speech auditory layer, sustaining 120ms per-frame inference on a Snapdragon 7 Gen 1 and throttling to 5 FPS to stay battery-neutral on mid-tier devices.

DeepS&P: Advanced LSTM-Based S&P 500 Quantitative Forecasting Engine | TensorFlow, Python *Dec 2025 – Feb 2026*

- Architected an institutional-grade AI forecasting platform for the S&P 500, deploying a 3-layer LSTM (256 hidden units, dropout-regularized) trained on 90+ years of market data.
- Developed quantitative analytics modules including interactive LSTM prediction validation and Monte Carlo simulations (2,000 stochastic paths) for volatility and scenario analysis.

GiftxAI: Retrieval-Augmented Recommendation Engine | LangChain, FAISS, Llama 3.3 70B, Streamlit *Dec 2025 – Aug 2026*

- Built an end-to-end RAG pipeline over unstructured PDF catalogs, chunking sources at 800 characters with 150-character overlap and indexing 384-dimensional all-MiniLM-L6-v2 embeddings in a FAISS store serving 740ms p50 query latency.
- Engineered an adaptive retrieval controller that expands context 5x on ranked queries (8 to 40+ chunks), paired with near-deterministic decoding (temperature 0.01) and numeric re-ranking, reaching 0.82 Relevance@5 on retrieved candidates.
- Implemented a configurable 3-pass validation chain over a 70B-parameter Groq-hosted Llama 3.3 backbone, enforcing source-grounded answers with per-response chunk citations and recording 0 hallucinations across a 200-query evaluation set.

QuantMaven: Multi-Asset Equity & Macro Research Terminal | Python, Streamlit, FRED API *Nov 2024 – Aug 2026*

- Developed a four-module research terminal (Dashboard, Leaderboard, Market, Economy) unifying 2 live sources, yfinance and the Federal Reserve FRED API, behind a single cached data layer supporting unlimited tickers.
- Implemented a technical analytics engine computing 6 indicators – 50/200-day SMAs, Bollinger Bands, 14-period RSI, realized volatility, and drawdown – across a 10-year lookback window on interactive candlestick and volume charts.
- Built a GICS leaderboard ranking bellwether tickers across all 11 sectors by mean daily return, period return, or price, alongside 4 macroeconomic series (GDP, fed funds rate, CPI, unemployment) and deep-linkable URL state for reproducible research views.

XGE: Explainable Gradient-Boosted Engine for Vehicle CO₂ Prediction | XGBoost, Scikit-learn *Sept 2024 – Nov 2024*

- Engineered an end-to-end regression pipeline predicting vehicle CO₂ emissions across 8 engineered features, chaining categorical label encoding and RobustScaler normalization into a 256-combination GridSearchCV sweep totalling 1,280 five-fold model fits.
- Tuned the XGBoost regressor to an R² of 0.95 at 9.0 g CO₂/km mean absolute error and 13.0 g/km RMSE, roughly 3% error against a 296 g/km fleet average, validated on a held-out 20% test split.
- Delivered model transparency via gain-based feature-importance attribution and residual diagnostics, deployed as a public Streamlit and Plotly dashboard for live what-if inference on user-supplied vehicle specifications.

Customer Churn Intelligence: Telecom Retention Classifier | Scikit-learn, SVM, KNN, PCA *Aug 2024 – Oct 2024*

- Built a supervised churn-prediction pipeline on telecom subscriber data, running full EDA, cleaning, encoding, and feature engineering before benchmarking Support Vector Machine and K-Nearest Neighbors classifiers.
- Tuned the SVM to 96.9% accuracy with a 0.90 F1 score, beating the KNN baseline by 1.3 points of accuracy and 0.05 F1 while lifting recall on the minority churn class from 0.76 to 0.85, a 12% gain in at-risk customers correctly flagged.
- Validated class separability with PCA projection and learning-curve diagnostics, confirming clean decision boundaries and stable generalization with no train/validation divergence.

EDUCATION

De La Salle University – Dasmariñas, B.S. in Computer Science

Aug 2022 – Aug 2026

GPA: 3.83 / 4.0, Consistent Dean's Lister & Loyalty Awardee

Leadership: CSITPC Director for Programming & Creatives; DLSUD CSO Director for Finance & Public Relations

CERTIFICATES

Ambient Agents with LangGraph - LangChain

July 2026

International Conference on Artificial Intelligence (ICAI 2026)

Feb 2026

Deep Learning Specialization – DeepLearning.AI (by Andrew Ng, Founder of Google Brain)

Dec 2025

- Neural Networks and Deep Learning
- Improving Deep Neural Networks
- Structuring Machine Learning Projects
- Convolutional Neural Networks
- Sequence Models

Data Science Workshop (by Dr. Jennifer Widom, Dean of Engineering, Stanford University)

Sept 2024

Data Analytics & Visualization Job Simulation – Accenture (Forage)

Feb 2024

Data Analysis with Python – freeCodeCamp

Nov 2023

Data Analytics using Excel – Great Learning

Aug 2023

Introduction to Analytics – Great Learning

May 2023

Google Analytics Certification – Google

Feb 2023